

## Approximation algorithms for the Geometric Covering Salesman Problem<sup>☆</sup>

Esther M. Arkin<sup>a,\*</sup>, Refael Hassin<sup>b,1</sup>

<sup>a</sup> *Department of Applied Mathematics and Statistics, Suny Stony Brook, Stony Brook, NY 11794-3600, USA*

<sup>b</sup> *Department of Statistics and Operations Research, School of Mathematical Sciences,  
Tel-Aviv University, Tel-Aviv 69978, Israel*

Received 16 August 1992; revised 2 June 1993

---

### Abstract

We introduce a geometric version of the Covering Salesman Problem: Each of the  $n$  salesman's clients specifies a *neighborhood* in which they are willing to meet the salesman. Identifying a tour of minimum length that visits all neighborhoods is an NP-hard problem, since it is a generalization of the Traveling Salesman Problem. We present simple heuristic procedures for constructing tours, for a variety of neighborhood types, whose length is guaranteed to be within a constant factor of the length of an optimal tour. The neighborhoods we consider include parallel unit segments, translates of a polygonal region, and circles.

---

### 1. Introduction

A salesman wants to meet a set of potential buyers. Each buyer specifies a compact set in the plane, his *neighborhood*, within which he is willing to meet. For example, the neighborhoods may be disks centered at the buyers locations, and each disk's radius specifies the distance that a buyer is willing to travel to the meeting place. The salesman wants to compute a tour of shortest length that intersects all of the buyers neighborhoods and finally returns to his initial departure point. (Note that the neighborhoods may overlap partially.) The problem generalizes the Euclidean Traveling Salesman Problem (TSP) in which the areas specified by the buyers are single points, and consequently it is NP-hard [16, 13].

On the other hand, it is known that the optimal tour of Euclidean Traveling Salesman (and in fact any symmetric TSP obeying the triangle inequality) can be

---

<sup>☆</sup> Partially supported by NSF Grants DMS 8903304 and ECSE-8857642.

\* Corresponding author. Email: estie@ams.sunysb.edu.

<sup>1</sup> E-mail: hassin@math.tau.ac.il.

approximated by a tour of length at most one and a half times the optimal tour [6]. Such approximation algorithms are available also for some generalizations of the TSP (see [4, 5, 9–12, 14, 17]). In this paper we construct algorithms with a bounded error ratio for some important cases of the Traveling Salesman with Neighborhoods Problem.

If all the neighborhoods are translates of each other, we can think of our problem as a “sweeper” problem: Given a broom of some shape, and points in the plane over which we wish to sweep, find the shortest path that will sweep over all required points. A continuous version of this problem, in which the points to be swept form a simple polygon, or a polygon with holes, is known as the *milling problem* and has a variety of applications (see [15, 3]).

The general method we use is to “represent” each neighborhood by a carefully chosen point in the neighborhood, and then apply a known approximation algorithm to these points in the plane. However, some naive choices for such representing points fail to deliver an approximation algorithm with a bounded error ratio. In Section 5, we discuss such examples. In Section 2, we give a method for choosing representative points for neighborhoods that are parallel unit segments. We show that this method does produce a constant approximation to the optimal tour. In Section 3, we describe some simple lower bounds on the length of the optimal tour. In Section 4, we discuss some extensions of this method to neighborhoods that are translates of a connected region. We also give a *Combination Lemma* that allows us to approximate a problem with regions of several different types, by combining approximations of each type. Thus for instance, we can approximate regions that are unequal length parallel segments, or segments parallel to one of  $k$  different directions.

We will assume (unless otherwise stated) that the initial location of the salesman can be viewed as a region of the same type as the customers regions. An alternative is to consider the salesman’s initial location as a point region and combine this region with an approximate tour on all other regions using the Combination Lemma.

It is interesting to compare our methods to those used by Current and Schilling [8]. The problem considered in their paper is a graph version of ours: Given a directed graph, non-negative costs associated with each arc, and a constant  $S$ , find a tour of minimum length such that all nodes not in the tour are at distance at most  $S$  from some node in the tour. Their heuristic proceeds by first finding a minimum vertex cover of the nodes and then approximating the shortest tour on the covering nodes. Unfortunately the first step of this procedure requires a solution of another NP-hard problem, and even if somehow this solution is obtained, there is no guarantee on how well this heuristic will perform. Our heuristic also starts with a covering problem which can be solved optimally in linear time after sorting, and results in a bounded performance ratio.

## 2. Parallel unit segments

In this section, we assume throughout that the regions are unit segments parallel to the  $x$ -axis. Let  $p$  denote the constant factor by which we can approximate an optimal

tour on a set of points in the plane. (Currently  $p = 1.5$ , [6].) Our result is the following theorem.

**Theorem 1.** *Given parallel equal length segments in the plane, we can find, in polynomial time, a tour visiting all segments, of length at most  $(3\sqrt{2} + 1)p$  times the length of an optimal such tour.*

**Proof.** Our approximation algorithm is simple: We first cover the unit segments by a minimum number of vertical lines. (A set of lines is said to *cover* a set of segments if each segment is intersected by at least one line from the covering set. We refer to the lines as *stabbers* or *covering lines*.) We do this in a greedy fashion. Our leftmost line is as far right as possible, namely at the leftmost right endpoint of a segment. Removing all segments covered by previous lines, we repeat this procedure, until all segments are covered. If one or two covering lines suffice, our approximation is trivial, and is described below. Otherwise, three or more covering lines are necessary, and the second step of our algorithm is to represent each unit segment by the point in which it intersects the covering lines. Note that by our construction of covering lines, each unit segment has a unique representative point. Finally we use these points as input to a bounded error TSP algorithm for points. Clearly, the resulting tour is a tour on the original segments. We will show that its length is within a constant factor of an optimal tour, but first we complete the discussion of the one or two covering lines cases.

It is interesting to note that we do not use the fact that the segments are of equal length, in the special case that one or two covering lines suffice. Indeed, as long as arbitrary length segments parallel to the  $x$ -axis can be stabbed by at most two lines parallel to the  $y$ -axis, an approximation algorithm is trivial. We use the following notation:  $y_1$  is the minimum  $y$ -value of a segment, and  $y_2$  the maximum value.

*Case 1* (All segments stabbed by a single line): It is easy to construct an optimal tour: Double the segment on a single covering line from  $y_1$  to  $y_2$ .

*Case 2* (Two covering lines are necessary and sufficient): We construct a tour as follows: Let  $x_1$  be the smallest  $x$ -value of a right endpoint of a segment. Let  $x_2$  be the greatest  $x$ -value of a left endpoint of a segment. Clearly,  $x_1 < x_2$ , otherwise one line could have covered all segments. Let  $y_1$  and  $y_2$  be as before. The tour constructed is a rectangle whose sides are parallel to the axes, cornered at  $(x_1, y_1)$ ,  $(x_2, y_1)$ ,  $(x_2, y_2)$ , and  $(x_1, y_2)$ . Clearly, this tour visits all segments, and its length is  $2(x_2 - x_1 + y_2 - y_1)$ . In Section 3 we show that  $LB_1 = 2\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$  is a lower bound on the length of an optimal tour. Hence we produced a tour of length at most  $\sqrt{2}$  times the length of an optimal tour.

*Case 3* (Three or more lines are needed to cover the segments): We will show that an optimal tour using the representative points is of length at most  $3\sqrt{2} + 1$  times the length of an optimal tour on the unit segments. We need some notation: Let  $l_j$ ,  $j = 1, \dots, k$  be the collection of covering lines given in increasing order of their  $x$  value. Let  $OPT$  be an optimal tour on the unit segments. Let  $\{d_i\}$  be the sequence of

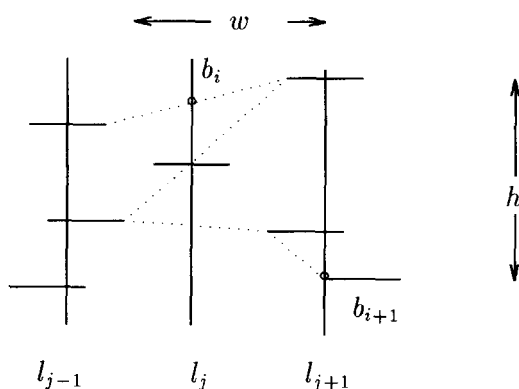


Fig. 1. Definition of a block.

intersection points of *OPT* with the lines, in cyclic order around *OPT*. We consider the corresponding sequence of intersected lines, and see that there may be multiple consecutive crossings of each particular line. We pick a subsequence of  $\{d_i\}$  corresponding to the first crossing point in each such consecutive sequence. To avoid double subscripts we denote this subsequence of the intersection points, in cyclic order around the tour, by  $\{b_i\}$ ,  $i = 1, \dots, m$ . Now, an optimal tour can be partitioned by the points  $\{b_i\}$  into blocks,  $B_i$ , which are the parts of the tour between  $b_i$  and  $b_{i+1 \pmod m}$ . (To simplify notation, from now on we drop the  $\pmod m$  term.)

Fig. 1 illustrates the definition of a block in which three stabbing lines are involved:  $l_{j-1}$ ,  $l_j$  and  $l_{j+1}$ . The optimal tour (shown in part by the dotted polygonal line) crosses line  $l_j$  three consecutive times, the first such crossing is defined to be  $b_i$ . Then *OPT* crosses  $l_{j+1}$ , and this crossing point defines  $b_{i+1}$ . The block  $B_i$  is the part of *OPT* from  $b_i$  to  $b_{i+1}$ . We define the *height* of a block  $B_i$ ,

$$h(B_i) \equiv \max_{(x,y) \in B_i} y - \min_{(x,y) \in B_i} y$$

where  $(x, y)$  are the coordinates of points in the block (and hence by definition points in the optimal tour). Similarly we define the *width* of a block,

$$w(B_i) \equiv \max_{(x,y) \in B_i} x - \min_{(x,y) \in B_i} x.$$

See Fig. 1. From these definitions it is easy to see that the length of a block  $B_i$  is at least  $\sqrt{h^2(B_i) + w^2(B_i)}$ . (This fact is formally proved in Section 3.) We use it to justify inequality (5) below. We drop the  $B_i$  when it is clear which is the block in question. Note that the width of  $B_i$  is at most one less than the horizontal distance between  $l_{j-1}$  and  $l_{j+1}$ . (This is true because a block starts and ends at two different covering lines, hence, either the rightmost or leftmost point of the block is a covering line. Thus we underestimate only on one side, and this underestimation is by at most the length of a segment.)

To complete the proof we exhibit an Eulerian graph,  $T$ , which is a union of subtours  $\{T_i\}_{i=1}^m$ , where  $T_i$  contains both  $b_i$  and  $b_{i+1}$ . These subtours satisfy two properties: First, each  $T_i$  “visits” all unit segments that  $B_i$  “visits”, but does so at the representative points of the segments (i.e., their crossing points with the covering lines). Second, the length of  $T_i$  is at most  $3\sqrt{2} + 1$  times the length of  $B_i$ . Since the length of an optimal tour on the representative points is no longer than the length of  $T$ , and the length of an optimal tour on the unit segments is the sum of the lengths of  $B_i$ , we get the desired result.

Denote  $I(B_i)$  the set of unit segments intersected by block  $B_i$ . The construction of  $T_i$  depends on  $I(B_i)$ , and on the lines covering these segments. We note that at most three lines cover the segments  $I(B_i)$ : the two (different) lines on which  $b_i$  and  $b_{i+1}$  lie (i.e., where the block begins and ends) and possibly the line on  $b_i$ 's opposite side. Without loss of generality, let  $b_i$  be on line  $l_j$  and  $b_{i+1}$  be on line  $l_{j+1}$ . The unit segments in  $I(B_i)$  are of three types: Segments covered by  $l_{j+1}$ , segments covered by  $l_j$ , and segments covered by  $l_{j-1}$ . By our construction, no other unit segment can be intersected by  $B_i$ . Of course,  $I(B_i)$  need not contain segments of all the three types and in particular it may even be an empty set.

For each line intersecting at least one unit segment in the block, its *top* (resp. *bottom*) with respect to the block, is the point on it intersected by the highest (resp. lowest)  $y$  valued unit segment in the block. If  $l_j$  ( $l_{j+1}$ ) does not intersect a segment, its top and bottom are both defined to be the  $y$  value of  $b_i$  ( $b_{i+1}$ ). We are now ready to describe  $T_i$ . There are two cases: (1)  $B_i$  visits no segments covered by line  $l_{j-1}$ , and (2)  $B_i$  visits at least one segment covered by  $l_{j-1}$ . In the first case,  $T_i$  is comprised of the following line segments: From the top of line  $l_j$ , to the bottom of the same line, to the bottom of the line  $l_{j+1}$ , to the top of that line, to the top of  $l_j$ . By Fact 1 the ratio of the length of  $T_i$  to the length of  $B_i$  is bounded by  $2\sqrt{2}$ .

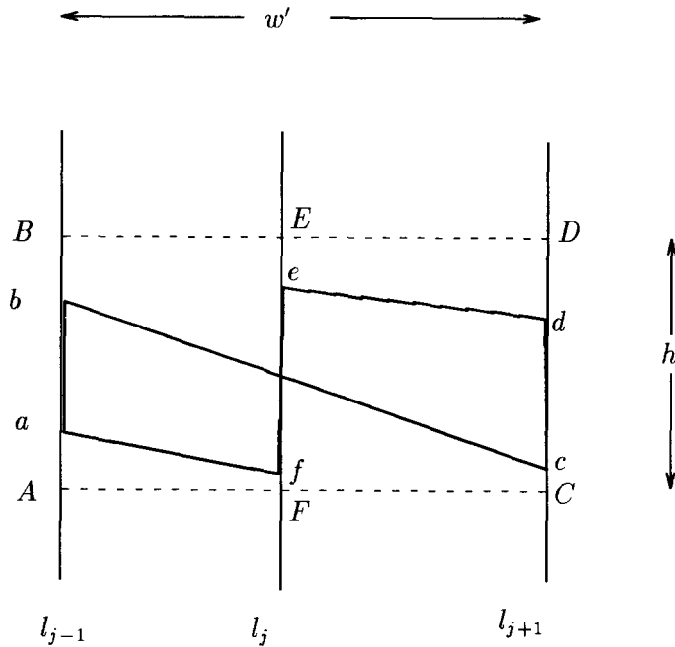
In the second case,  $T_i$  is comprised of the following line segments: From the top of line  $l_{j-1}$ , to the bottom of the same line, to the bottom of the line  $l_j$ , to the top of that line, to the top of  $l_{j+1}$ , to the bottom of that line, to the top of  $l_{j-1}$ . We claim that ratio of the length of  $T_i$  to the length of  $B_i$  (which we denote by  $|B_i|$ ) is bounded by  $3\sqrt{2} + 1$ .

To see this, consider the rectangle of size  $hw'$  that encloses  $B_i$ , with corners  $A, B, C, D$  in clockwise order from the lower left, where  $A, B$  are on  $l_{j-1}$ ,  $C, D$  on  $l_{j+1}$ . Let  $E$  on its upper edge,  $F$  on its lower edge, be the points on  $l_j$ . Note that  $w' \leq w + 1$ . (See Fig. 2.) We are looking for the positions of points  $a, b, c, d, e, f$  (where  $a, f, c$  are above  $A, F, C$  and  $b, e, d$  below  $B, E, D$  but above  $a, f, c$  as in the figure) that will maximize the ratio of the length of the tour  $a, b, c, d, e, f, a$  to the diagonal of the rectangle that encloses  $B_i$ , which is  $\sqrt{h^2 + w^2}$ .

Clearly, such a tour will have  $b = B$  and  $c = C$ . Then, if  $a$  has a smaller  $y$  coordinate than  $f$  the tour will have  $a = A$  and then  $f = F$ . Else, it will clearly have  $f = F$  and then also  $a = A$ . Similarly, the tour will have  $E = e$  and  $D = d$ . The length of the tour is then

$$|T_i| \leq (w + 1) + 3h + \sqrt{h^2 + (w + 1)^2} \quad (1)$$

$$\leq (w + 1) + 3h + \sqrt{h^2 + w^2} + 1 \quad (2)$$

Fig. 2. The tour  $T_i$ .

$$\leq 3(w + h) + \sqrt{h^2 + w^2} \quad (3)$$

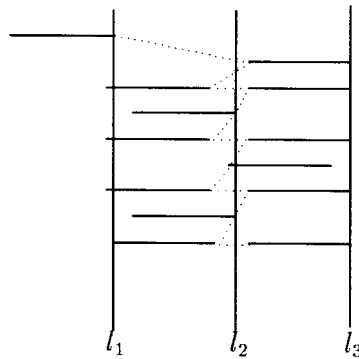
$$\leq (3\sqrt{2} + 1)\sqrt{h^2 + w^2} \quad (4)$$

$$= (3\sqrt{2} + 1)|B_i|. \quad (5)$$

Inequality (1) follows by the triangle inequality. Inequality (3) relies on the fact that  $w \geq 1$ . This is true because the width of any block is at least the distance between two consecutive covering lines, which is at least one, the length of the segments. Inequality (4) follows from the fact that  $(w + h) \leq \sqrt{2}\sqrt{w^2 + h^2}$ . In Section 3 we show that a closed walk touching all four sides of a rectangle has length at least twice the diagonal. For (5) we are using a consequence of this fact which implies that a walk that is not necessarily closed but visits all four sides of a rectangle has length at least the diagonal of the rectangle.

Notice that we have the vertical portions of  $T_i$  traverse each of the covering lines, each line between its top and bottom, and thus  $T_i$  visits all unit segments visited by  $B_i$ . Hence  $T$  is an Eulerian graph meeting all segments, in an order possibly different from the order they are visited by the optimal tour. This concludes our proof for parallel unit line segments.  $\square$

It is interesting to note that the approximate tour we obtain may visit the segments in a different order than an optimal tour. Fig. 3 shows a partial example in which the

Fig. 3. *APX* uses a different order than *OPT*.

order in which the approximate tour visits the segments is very different from the order used by the optimal tour. (The figure is exaggerated, and should be understood to imply that the segments are all very close to the covering line  $l_2$ . The partial tour is shown by a dotted line.) The part of the optimal tour shown here is all a single block. The order *APX* uses is to first visit all segments stabbed by  $l_1$ , then all stabbed by  $l_2$  and last the segments stabbed by  $l_3$ .

### 3. Lower bounds

Our first lower bound,  $LB_1$ , is derived by considering a rectangle for which we know that the optimal tour (and in fact any tour visiting all regions) must “touch” all of its four sides. We begin with regions that are unit segments parallel to the  $x$ -axis.

Let  $x_1$  be the smallest  $x$ -value of a right endpoint of a segment. Let  $x_2$  be the greatest  $x$ -value of a left endpoint of a segment. Assume  $x_1 \leq x_2$ . (If this assumption is not satisfied then the segments have a common  $x$  coordinate and an optimal solution is obvious.) Let  $y_2$  be the maximum  $y$ -value of a segment,  $y_1$  the minimum value. Set  $LB_1 = 2\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ . Any tour visiting all segments must go as far to the left as  $x_1$ , as far to the right as  $x_2$ , as far down as  $y_1$  and as far up as  $y_2$ . Thus a tour must visit all four sides of the rectangle whose corners are  $(x_1, y_1)$ ,  $(x_1, y_2)$ ,  $(x_2, y_2)$ , and  $(x_2, y_1)$ . The following fact which we prove below shows that any tour that touches all four sides of a rectangle, must have length at least twice the diagonal, implying that  $LB_1$  is a lower bound for *OPT*. If we wish to include the special instances for which  $x_1 > x_2$  in this lower bound we can write more generally  $LB_1 = \sqrt{((x_2 - x_1)^+)^2 + (y_2 - y_1)^2}$ . (We use the notation  $a^+ \equiv \max(0, a)$ .)

**Fact 1.** *A tour touching all four sides of a rectangle is of length at least twice the diagonal of the rectangle.*

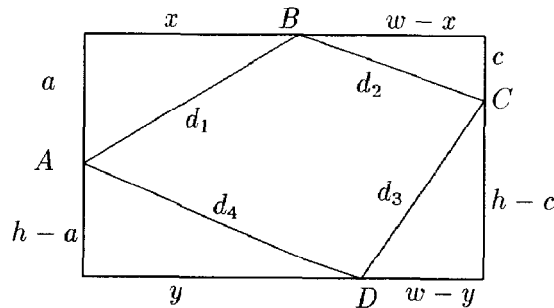


Fig. 4. A tour touching four sides of a rectangle.

**Proof.** Consider a shortest tour touching all four sides of the rectangle. Clearly, such a tour is non-self intersecting. Thus we may think of this tour as starting at a point  $A$  on the left vertical side of the rectangle, continue to  $B$ , a point on the top horizontal side of the rectangle, to  $C$ , a point on the right vertical side, to  $D$  a point on the bottom side and back to  $A$ . Let  $d_1$  be the length of the interval between  $A$  and  $B$ ,  $d_2$ ,  $d_3$ , and  $d_4$  the interval lengths of the segments  $BC$ ,  $CD$ , and  $DA$  respectively. Let  $d$  be the length of the diagonal of the rectangle, which is at height  $h$  and width  $w$  (see Fig. 4).

We observe that the angle at which the tour hits each side is equal to the angle with which it departs that side, by Snell's law.

Consider the four right triangles created by an edge of the tour and the rectangle. These triangles are similar since their angles are the same. Let  $x$ ,  $w - x$ ,  $a$  and  $c$  be the length as before, and let  $h - a$  (resp.  $h - c$ ) be the length of the segment between the bottom left (resp. right) corner and  $A$  (resp.  $C$ ). Let  $y$  and  $w - y$  be the lengths of the segments between the bottom left corner and  $D$ , and the bottom right corner and  $D$  respectively. By the similarity we have

$$\frac{x}{a} = \frac{y}{h-a} = \frac{w-y}{h-c} = \frac{w-x}{c}.$$

Thus  $y = (h-a)(x/a)$ ,  $h-c = (w-(h-a)(x/a))(a/x)$ , and  $c = h - (w(a/x) - (h-a))$ . Using similarity again, we get that the last length  $c$  is also equal to  $(a/x)(w-x)$ . Cleaning up terms we get  $x/a = w/h$ . Now  $d_1 = x/\sin \alpha$  and  $d_2 = (w-x)/\sin \alpha$  and  $\sin \alpha = w/d$ , yielding that  $d_1 + d_2 = d$  and similarly  $d_3 + d_4 = d$ .  $\square$

We can state a corresponding lower bound for more general regions. Let the diameter,  $\delta$ , of a region be the distance between the two points in the region farthest apart. We consider the case in which the diameters are parallel segments, such as when the regions are all translates of the same shape. Without loss of generality we assume that the diameter is between two points whose  $y$ -value is the same (i.e., the two points in the region determining the diameter lie parallel to the  $x$ -axis). Here, we define  $x_1, x_2, y_1$  and  $y_2$  as above, using the diameters of the regions as the segments. Next, let  $y_1(R)$  be the maximum  $y$ -value of the region by which  $y_1$  was defined, and let  $y_2(R)$  be



the minimum  $y$ -value of the region by which  $y_2$  was defined. We may have  $y_1(R) \geq y_2(R)$ . The lower bound on the length of the optimal tour is now  $LB_1 = 2\sqrt{((x_2 - x_1)^+)^2 + ((y_2(R) - y_1(R))^+)^2}$ .

In Section 4 we will use another rectangle to generate such a lower bound. Consider the smallest perimeter rectangle touching all regions. The fact that this rectangle has minimal perimeter implies that there are contact-critical points, one on each side of the rectangle, where a region “barely touches” the rectangle. Thus again we can lower bound the length of an optimal tour by twice the length of the diagonal of this rectangle. Note that we do not have to restrict such a rectangle to have sides parallel or perpendicular to the diameters, any direction will do. The important property is that an optimal tour must visit all four sides of the rectangle and thus have length bounded below by twice the length of its diagonal.

A second lower bound can be obtained by considering distances between pairs of regions. Let  $d_{ij}$  be the distance between regions  $i$  and  $j$ , measured as the distance between the nearest pair of points on these two regions. Consider a complete graph  $G$  where each node corresponds to a region and the length of the arc connecting nodes  $i$  and  $j$  is  $d_{ij}$ . Let  $LB_2$  be the length of a shortest tour on  $G$ . Clearly,  $LB_2$  is a lower bound on  $OPT$ .

Let  $LB = \max\{LB_1, LB_2\}$ . Fig. 5 demonstrates that  $LB/OPT$  may be arbitrarily close to zero, even when the regions are parallel equal length segments. In this example there are  $n$  segments, divided into  $\sqrt{n}$  “zigzags”, each of which contains  $\sqrt{n}$  segments. In this case we see that  $LB_1$  is determined by the dotted rectangle,  $LB_2$  is two times the height of the dotted rectangle. On the other hand, as  $n$  tends to infinity, the segments are very short compared to the sides of the rectangle, and in fact are almost like points densely spread in the rectangle. The length of an optimal tour on such segments tends to infinity as  $n$  tends to infinity.

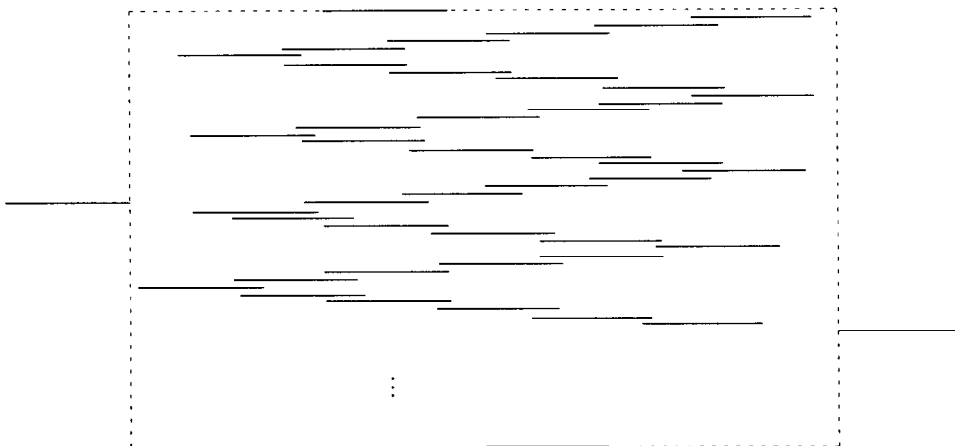


Fig. 5.  $LB/OPT$  approaches zero.

It is interesting to see that, in the above example,  $LB_2$  (and thus possibly  $LB$ ) may be *increased* if we *delete* regions, resulting in a tighter lower bound. This idea can be formalized as follows: Let  $S$  be the set of regions. For each subset  $S' \subseteq S$  let  $LB_2(S')$  be the resulting lower bound when all regions in  $S \setminus S'$  are deleted. Then set  $LB'_2 = \max_{S' \subseteq S} \{LB_2(S')\}$ , and  $LB' = \max\{LB_1, LB'_2\}$ . It is an open question whether the ratio  $LB'/OPT$  can be made arbitrarily close to zero or it is bounded below by some positive constant. If the latter is true, it is interesting to ask whether the optimizing  $S'$  can be computed efficiently (i.e., in polynomial time).

## 4. Extensions

### 4.1. Translate regions

Our next generalization is to regions that are translates of the same convex body, e.g., a unit circle or rectangle. (Simple modifications that allow the regions to be non-convex are discussed below.) Our idea is to imitate our algorithm for segments. Recall, the *diameter*,  $\delta$ , of a region is the distance between the two points in the region farthest apart. Without loss of generality we assume that the diameter is between two points whose  $y$ -value is the same. Now treating these diameters as equal parallel segments of length  $\delta$ , we find a minimum cover by vertical lines (covering these diameters). Next, we pick a representing point from each region to be the point of intersection of the diameter and the covering line. By the convexity assumption, this point is in the region. Let  $\eta$  be the height of a region, namely the vertical distance between the points in a region with highest and lowest  $y$ -value. Note that  $\eta \leq \delta$ . We further define  $\eta_1$  ( $\eta_2$ ) to be the vertical distance between the representing point and the highest (resp. lowest) point in the region. By definition  $\eta = \eta_1 + \eta_2$ . Again, we separate our discussion to the cases in which one, two, or three or more covering lines are necessary. However, unlike the unit segment case, in which the one and two covering lines cases were trivial, here, these are the more difficult to extend. The intuitive reason is that very short optimal tours are possible in these cases, and a constant factor approximation is harder to obtain.

A first attempt to extend treatment of the segment regions to convex (or general) regions if one line suffices to stab all regions is to pick again a vertical segment that is part of the stabbing line and double it. The following example illustrates the failure of this straightforward generalization.

**Example 1.** The regions are disks. Consider two unit disks separated by and tangent to a vertical line. Let  $2x$  be the length of the vertical segment between the tangent points, and hence the approximation is  $4x$ . Let  $2y$  be the distance between the disks (implying that  $4y$  is the optimal tour length). Then  $(1 + y)^2 = 1 + x^2$ , so the ratio of the approximated tour to the optimal tour is equal to  $x/y = \sqrt{1 + (2/y)^2}$ , which tends to infinity as  $y$  tends to zero (see Fig. 6).

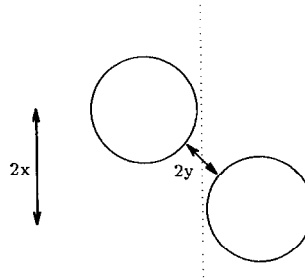


Fig. 6. Example 1.

Next we describe an approximation method for translate convex regions which does produce a constant performance ratio. Recall that  $p$  denotes the constant factor by which we can approximate an optimal tour on a set of points in the plane. We need the following simple fact.

**Fact 2.** For constants  $a$  and  $b$ , the following inequality holds for all  $w$  and  $h$ :

$$aw + bh \leq \sqrt{a^2 + b^2} \sqrt{w^2 + h^2}.$$

**Proof.** Immediate, by squaring both sides of the inequality.  $\square$

**Theorem 2.** Given translates of a convex region in the plane we can find in polynomial time, a tour visiting all regions, of length at most  $(\sqrt{7^2 + 3^2} + 1)p$  times the length of an optimal such tour.

**Proof.** We separate our discussion into three cases, depending on whether one two or more lines are necessary to cover the diameters of the regions.

*Case 1 (One stabbing line suffices):* Find the smallest perimeter rectangle, whose sides are aligned with the axes, that touches all regions. (Here, and whenever discussing minimum perimeter rectangles touching all regions, we consider a rectangle to be the two dimensional region enclosed by its perimeter.) Note that some regions may lie completely inside the rectangle. Denote the width of the rectangle by  $W$  and its height by  $H$ . By Fact 1, the length of an optimal tour is at least  $2\sqrt{W^2 + H^2}$ . Note that the perimeter of the rectangle may not be a “legal” tour for the regions, because it may not visit all regions (namely the regions completely inside the rectangle). We add (twice) the vertical segment from the bottom of the rectangle to its top. This doubled segment is placed at the middle of the horizontal sides. With this addition we get a tour that is guaranteed to visit all regions. This crucially relies on the fact that all regions can be stabbed by a single vertical line, and thus all lie in a vertical strip of width  $2\delta$ . Hence  $W \leq 2\delta$ , implying that the regions are at least as wide as half of the rectangle. The

length of this tour is  $2W + 4H$  which is, by Fact 2, at most  $2\sqrt{5}\sqrt{W^2 + H^2}$  which is at most  $\sqrt{5}$  times the optimal tour length.

We discuss briefly how to find (in polynomial time) a minimum-perimeter rectangle touching all regions whose sides are parallel to the coordinate axes. If the regions in question are simple polygons, then a minimum-perimeter rectangle is determined by four contact points, which will be vertices of the regions touching the edges of the rectangle. A naive algorithm follows immediately: Examine all rectangles determined by quadruples of region vertices and edges, check each to see if it touches all regions, and select a minimum-perimeter such rectangle. If the regions are circles, then it is easy to show that the only contact points between circle boundaries and the boundary of a minimum-perimeter rectangle are points tangent to lines parallel to the axes and to  $45^\circ$  lines (in order to accommodate corner solutions). The naive algorithm can then be applied to this case as well. Using techniques similar to [1], a faster algorithm can be designed (private communication with Suri). For a set of convex polygons, an  $O(n)$  time algorithm using linear programming with four variables, has been obtained by D. Rappaport (private communication).

*Case 2 (Two stabbing lines):* Let us assume that we pick the two stabbing lines to be as close to each other as possible, and denote by  $D$  the (horizontal) distance between these two lines. There are two cases to consider: 2.1.  $D \geq \delta$ , and 2.2.  $D < \delta$ .

*Case 2.1 ( $D \geq \delta$ ):* This case is similar to the unit segment case. Define  $x_1$  to be the smallest  $x$ -value of a right endpoint of a diameter, and  $x_2$  be the largest  $x$ -value of a left endpoint of a diameter (as we did for the unit segment case). Then by definition  $x_2 - x_1 = D$ . Next, let  $y_1$  be the minimum  $y$ -value of a diameter, and  $y_1(R)$  be the maximum  $y$ -value of that region. Similarly, let  $y_2$  be the maximum  $y$ -value of a diameter and  $y_2(R)$  be the minimum  $y$ -value of that region. Again these correspond to the unit segment case, where  $y_1 \leq y_2$ , but we may have  $y_1(R) \geq y_2(R)$ . However,  $y_2 - y_1 \leq (y_2(R) - y_1(R))^+ + \eta$ , where  $\eta$  is the height of the regions and  $a^+ \equiv \max(a, 0)$ .

A lower bound on the length of the optimal tour is  $2\sqrt{(x_2 - x_1)^2 + ((y_2(R) - y_1(R))^+)^2}$  (see Section 3). The approximation tour we construct is a rectangle whose sides are parallel to the axes, and cornered at  $(x_1, y_1)$ ,  $(x_2, y_1)$ ,  $(x_2, y_2)$ , and  $(x_1, y_2)$ . Let the length of this tour be denoted by  $APX$ , and the length of the optimal tour be denoted by  $OPT$ .

$$\begin{aligned} APX &= 2(x_2 - x_1) + 2(y_2 - y_1) \\ &\leq 2D + 2(y_2(R) - y_1(R))^+ + 2\eta \\ &\leq 4D + 2(y_2(R) - y_1(R))^+ \end{aligned} \tag{6}$$

$$\begin{aligned} &\leq 2\sqrt{5}\sqrt{D^2 + ((y_2(R) - y_1(R))^+)^2} \\ &\leq \sqrt{5}OPT. \end{aligned} \tag{7}$$

Here (6) follows from our assumption of Case 1:  $\eta \leq \delta \leq D$  and (7) follows from Fact 2.

*Case 2.2* ( $D < \delta$ ): Here we find a rectangle of minimum perimeter, whose sides are parallel to the axes, that touches all regions. Denote the width of the rectangle by  $W$  and its height by  $H$ . By Fact 1, the length of an optimal tour is at least  $2\sqrt{W^2 + H^2}$ . Note that the perimeter of the rectangle may not be a tour, because it may not visit all regions (namely the regions completely inside the rectangle). However, adding (twice) the vertical segments from the bottom of the rectangle to its top, at its one third and two third points width-wise, does produce a tour guaranteed to visit all regions. The reason is similar to the one stabbing line case: All regions are stabbed by two vertical lines which are separated by at most  $\delta$  (by the assumption  $D < \delta$ ), and thus all lie in a vertical strip of width  $3\delta$ . Hence  $W \leq 3\delta$ , implying that the regions are at least as wide as one third of the rectangle. The length of this tour is  $2W + 6H$ , which is at most  $2\sqrt{10}\sqrt{W^2 + H^2}$ , which is at most  $\sqrt{10}$  times the optimal tour length.

*Case 3* (Three or more covering lines): This case is very similar to the segment case. We define blocks,  $B_i$ , and their heights and widths as before. The definition of the top (and bottom) of a line with respect to a block is modified slightly to reflect the fact that regions have a height. The top of the line in a block is the higher of the highest point visited by the block and the highest representing point in the block. Similarly define the bottom of a line with respect to a block. Noting that the top (bottom) of a line with respect to block  $B_i$  is at most  $\eta_2$  higher ( $\eta_1$  lower) than the highest (lowest) crossing point of this line by block  $B_i$ , we get that the distance between the top and bottom of a line in block  $B_i$  is at most  $\eta + h(B_i)$ .  $T_i$  is defined as before noting the modifications of the top and bottom definitions. Clearly,  $T_i$  visits all regions that are visited by  $B_i$ . To bound the length of  $T_i$  we have, instead of Equations (1)–(5):

$$|T_i| \leq (w + \delta) + 3(h + \eta) + \sqrt{(h + \eta)^2 + (w + \delta)^2} \quad (8)$$

$$\leq w + \delta + 3(h + \eta) + |B_i| + \delta + \eta \quad (9)$$

$$\leq 7w + 3h + |B_i| \quad (10)$$

$$\leq (\sqrt{7^2 + 3^2} + 1)|B_i|. \quad (11)$$

Inequality (10) relies on the fact that  $\eta \leq \delta$ , and  $w(B_i) \geq \delta$ . Inequality (11) follows from Fact 2. This concludes our proof for translates of a convex region.  $\square$

The analysis for the case of three or more covering lines did not require the full description of the body, of which the regions were translates, only its diameter. In fact we do not require that all regions be translates of one body; it suffices that the diameters of all the regions are parallel equal length segments, and that the regions are convex. However, the seemingly simpler cases in which one or two covering lines suffice require us to find a rectangle as described. This can be done in polynomial time for regions such as polygons or splinegons, but may present a problem for more general regions.

Translates of connected non-convex regions can also be approximated, assuming their representation is such that the computation of the minimum perimeter rectangle

is easy. In the following theorem, the diameter of a region is the maximum Euclidean distance between any pair of its points.

**Theorem 3.** *Given translates of a connected (not necessarily convex) region in the plane we can find, in polynomial time, a tour visiting all regions, of length at most  $(\sqrt{11^2 + 3^2} + 1)p$  times the length of an optimal such tour.*

**Proof.** We begin, as in other cases, by covering the regions greedily by vertical lines. Since the regions are connected, we have, as before, that the distance between covering lines is at least the diameter of the regions. (Note that if the regions are not connected, a greedy cover might use very close stabbers. As a result, we are not able to use an inequality similar to (2) to prove a bound in this case.)

If one or two covering lines suffice to cover the regions, then our approximation scheme is identical to the convex case.

If three or more vertical lines are necessary to cover the diameters, only a slight modification to the definitions and analysis is needed to obtain a constant error ratio. The ratio obtained is only somewhat worse than the convex case. We must modify our definition of a representing point, since the intersection between the covering lines and diameter of a region may be outside a region. Instead, we choose as a representing point (arbitrarily) any point in the intersection of the region with the covering line. The top (bottom) of a line with respect to a block is the point on the line with highest (resp. lowest)  $y$ -value in a region visited by this block. Here we bound the top (and bottom) of a line in a block to be at most  $\eta$  away from the highest (lowest) point visited by the block, and so the distance between the top and bottom of a block is at most  $2\eta + h(B_i)$ . We proceed as before:

$$\begin{aligned} |T_i| &\leq (w + \delta) + 3(h + 2\eta) + \sqrt{(h + 2\eta)^2 + (w + \delta)^2} \\ &\leq w + \delta + 3(h + 2\eta) + |B_i| + \delta + 2\eta \\ &\leq 11w + 3h + |B_i| \\ &\leq (\sqrt{11^2 + 3^2} + 1)|B_i|. \quad \square \end{aligned}$$

#### 4.2. Combining approximations

The lemma we describe next, which we refer to as the *Combination Lemma*, allows us to approximate a problem with regions of several different types, by combining approximations of each type. This lemma can be applied for instance, to the case in which the regions are unit segments parallel to one of  $k$  different directions (e.g.,  $k = 2$  and segments are parallel to either the  $x$ -axis or the  $y$ -axis). The error ratio obtained is  $k(c + 2) - 2$ , where  $c$  denotes the error ratio of the single direction problem. Another application is to the case in which the segments are parallel, but may be of one of  $k$  different lengths, including zero length segments, namely points.

**Lemma 1** (Combination Lemma). *Given regions that can be partitioned into two types, and constants  $c_1, c_2$  bounding the error ratios which we can approximate the optimal tours on regions of types 1 and 2, then we can approximate the optimal tour on all regions with an error ratio bounded by  $c_1 + c_2 + 2$ .*

**Proof.** Let  $OPT_1$  and  $OPT_2$  be the optimal tour lengths for regions of type 1 and 2. Let  $OPT$  be the overall optimal tour length. By the triangle inequality, each subproblem's optimal value is bounded by the optimal value to the original problem ( $OPT_i \leq OPT$ ). Denote by  $APX_1, APX_2$  and  $APX$  the approximate tour lengths for regions of type 1, 2, and all regions, obtained by methods described below. We now describe two heuristics. Our algorithm constructs the two approximations produced by these heuristics, and chooses the best of them. We will distinguish two cases, and describe for each of them the heuristic that guarantees the claimed bound. Let  $\delta_1$  and  $\delta_2$  be the diameters of the two region types. Our proof consists of two cases: Case 1:  $2\delta_1 + 2\delta_2 \leq OPT$ , Case 2:  $2\delta_1 + 2\delta_2 > OPT$ .

*Case 1:* Obtain  $APX_1$  and  $APX_2$  by the hypothesis of the theorem, with corresponding bounds  $c_1$  and  $c_2$ . Let  $D$  be the minimum distance between a point in a type 1 region and a point in a type 2 region. Clearly  $2D \leq OPT$ . We obtain  $APX$  by combining the two approximate solutions into a tour visiting all regions by “gluing” the tours together at the place in which the two region types are closest to each other. This “glue” has length bounded by  $2(D + \delta_1 + \delta_2)$ . Thus we have

$$\begin{aligned} APX &\leq APX_1 + APX_2 + 2D + 2\delta_1 + 2\delta_2 \\ &\leq c_1 OPT_1 + c_2 OPT_2 + OPT + OPT \\ &\leq (c_1 + c_2 + 2)OPT. \end{aligned}$$

*Case 2:* We begin by constructing a minimum perimeter rectangle (whose sides are parallel or perpendicular to a fixed direction of our choice) that touches all regions (of both types). Denote the lengths of the sides of the rectangle by  $W$  and  $H$ . We know that  $OPT \geq 2\sqrt{W^2 + H^2}$ . Without loss of generality we assume that  $\delta_1 \leq \delta_2$ . We further partition Case 2 into two subcases: Case 2a:  $\delta_1 \geq OPT/4$  (and the definition of Case 2, again  $\delta_2 > OPT/4$ ), Case 2b:  $\delta_1 < OPT/4$  (and hence  $\delta_2 > OPT/4$ ).

*Case 2a:* Build  $APX$  by going around the perimeter of the rectangle combined with two (doubled) stabbing segments, one for each region type, that visit all regions not visited by the perimeter of the rectangle. Finding such stabbers is an easy task: We ignore all regions stabbed by the boundary of the rectangle and find a line cover for regions of each type, using lines perpendicular to the direction of the diameter. We claim that one line suffices, since the length of each diameter is at least half the length of the rectangle's diagonal in Case 2a. In fact, only the part of the covering line inside the rectangle is sufficient to stab all regions of one type completely in the rectangle. We use this segment of the covering line, as the stabbing segment needed by  $APX$ . The length of the segment is bounded by the length of the diagonal of the rectangle, and since we are adding two segments, each doubled, our approximation length is

bounded by the perimeter of the rectangle  $(2W + 2H)$  plus four diagonals  $(4\sqrt{W^2 + H^2})$ :

$$\begin{aligned} APX &\leq 2W + 2H + 4\sqrt{W^2 + H^2} \\ &\leq 8\sqrt{W^2 + H^2} \\ &\leq 4OPT. \end{aligned}$$

*Case 2b:* Obtain  $APX_1$  by the hypothesis of the theorem, with bound  $c_1$ . Build  $APX_2$  as in Case 2a, this time using only one stabber, to visit only regions of type 2 not visited by the rectangle perimeter. Next we glue together the two approximations. The length of this glue depends on the distance between the rectangle of  $APX_2$  and  $APX_1$ . If  $APX_1$  crosses the rectangle, clearly, the glue length is zero. If  $APX_1$  lies completely inside the rectangle then the length of glue is at most the minimum of  $W$  and  $H$ . The remaining possibility is that  $APX_1$  lies completely outside of the rectangle, but recall this rectangle touches all regions, and hence  $APX_1$  can be at most  $\delta_1$  away from the rectangle. In this case glue of length  $2\delta_1 < OPT/2$  suffices. In summary, in all three cases the glue length is bounded by  $OPT/2$ :

$$\begin{aligned} APX &\leq APX_1 + 2W + 2H + 2\sqrt{W^2 + H^2} + \text{glue} \\ &\leq c_1OPT_1 + 2(1 + \sqrt{2})\sqrt{W^2 + H^2} + OPT/2 \\ &\leq c_1OPT + (1 + \sqrt{2})OPT + OPT/2 \\ &\leq (c_1 + 1.5 + \sqrt{2})OPT. \end{aligned}$$

To complete the proof we recall that  $c_i \geq 1$  in all cases, and thus the bounds of Cases 2a and 2b are at least as good as the bound claimed in the lemma.  $\square$

We can use this lemma repeatedly to obtain approximations to more than two region types. The bound we obtain for combining  $k$  different regions types with individual approximation bounds of  $c_1, c_2, \dots, c_k$  is  $c_1 + c_2 + \dots + c_k + 2(k - 1)$ . In the section below, we show how to use the Combination Lemma to obtain an approximation in the case that the regions are parallel line segments of varying lengths.

#### 4.3. Unequal segments

We now describe two algorithms that approximate the optimal tour when the regions are parallel segments of arbitrary lengths, although point regions are not allowed. (Point regions can most easily be incorporated using the Combination Lemma.) Without loss of generality, let the shortest region be a segment of unit length, and the longest segment be of length  $r$ . To simplify matters, we further assume that  $r$  is an integer, otherwise we can replace  $r$  by its ceiling. The first algorithm and its analysis are straightforward generalizations of the previous ones: Cover the segments by



a minimum number of vertical lines. If one or two lines suffice to cover the segments, the analysis is identical to the equal length segment case. Otherwise, choose as a representative point of each segment, the rightmost intersection with a covering line. Define blocks, their heights and widths as before. Note that whereas a block visited segments covered by at most three different lines when all segments were of the same length, now a block may visit segments whose representative point is on one of at most  $r + 2$  different covering lines. This follows from our choice of the representative point as the rightmost crossing by a stabbing line. (Leftmost crossing would work equally well. However, if we were to choose any crossing point as the representative point, a block could possibly visit segments stabbed by up to  $2r + 1$  different stabbing lines.) Furthermore, the width of a block is at most  $2r$  less than the distance between the left and rightmost covering lines in the block, because we may be “off” by  $r$  on each side of the block. Note that we cannot improve this bound to  $2r - 1$  and thus match the bound for the equal case. To see this, let  $b_i$  be on covering line  $l_j$  and  $b_{i+1}$  on line  $l_{j+1}$ . The block can contain regions covered by lines  $l_{j-1}$  and  $l_{j+r}$  which are at distance at most  $r$  from the tour on each of its sides. We construct  $T_i$  to go between the top and bottom of each line in the block (if the line has a representative point of any visited segments on it). Between covering lines,  $T_i$  simply travels from the top (bottom) of one line to the top (bottom) of another, as necessary. The length of this vertical and horizontal traveling of subtour  $T_i$  is bounded by  $(r + 2)h(B_i) + w(B_i) + 2r + \sqrt{h^2 + (w(B_i) + 2r)^2}$ . We conclude the proof as before:

$$\begin{aligned}
 |T_i| &\leq (r + 2)h(B_i) + w(B_i) + 2r + \sqrt{h(B_i)^2 + (w(B_i) + 2r)^2} \\
 &\leq (r + 2)h(B_i) + (4r + 1)w(B_i) + \sqrt{h(B_i)^2 + w(B_i)^2} \\
 &\leq f(r)\sqrt{h^2(B_i) + w^2(B_i)} \\
 &\leq f(r)|B_i|.
 \end{aligned}$$

Here  $f(r) = \sqrt{(r + 2)^2 + (4r + 1)^2} + 1$  is obtained from Fact 2, and can be bounded above by the simpler expression  $\sqrt{17r} + \sqrt{5} + 1$ . If  $r$  is even, this bound can be improved slightly, since the closing of the subtour  $T_i$  will be bounded by  $w(B_i) + 2r$  instead of the diagonal  $\sqrt{h(B_i)^2 + (w(B_i) + 2r)^2}$ . The resulting bound in the even  $r$  case is thus  $f(r) = \sqrt{5}(r + 2)$ . Recall that  $p$  denotes the constant factor by which we can approximate an optimal tour on a set of points in the plane. We have shown the following theorem.

**Theorem 4.** *Given parallel segments in the plane, of lengths between 1 and  $r$ , we can find, in polynomial time, a tour visiting all segments, of length at most  $f(r)\sqrt{2p}$  times the length of an optimal such tour.*

Note that the bound is not quite as good as the one obtained when  $r = 1$ , namely all segments are of equal length. We conclude this section by describing a second algorithm, which uses the Combination Lemma to obtain an approximation bound of

$O(\log r)$  for regions that are parallel line segment of length between 1 and  $r \geq 2$ . First, divide the segments into  $\log r$  classes, where class  $i$  contains all segments of lengths between  $2^{i-1}$  and  $2^i$ . Now approximate the optimal tour on each class using the method above. Note that within each classes the ratio of the longest segment to the shortest is bounded by 2, so this yields an approximation factor of  $c_i = (3 \cdot 2 + 3)\sqrt{2} = 9\sqrt{2}$ . Using the Combination Lemma  $\log r$  times yields a bound of  $(9\sqrt{2} + 2) \log r$ . We have shown the following theorem.

**Theorem 5.** *Given parallel segments in the plane, of lengths between 1 and  $r \geq 2$ , we can find, in polynomial time, a tour visiting all segments, of length at most  $(9\sqrt{2} + 2)p \log r$  times the length of an optimal such tour.*

One might ask whether a better bound in the case of unequal length parallel segments is possible. In particular, it is possible to get a bound in which  $r$ , the ratio of the longest to shortest segment, does not appear. This does not seem possible using our algorithm as the example in Fig. 7 shows. In this example covering lines are determined by the short segments on top, and every longer segment is intersected by covering lines many times. As long as we restrict our choice of which such an intersection point we select to represent a segment to the rightmost, or leftmost, or “middle” intersection, the resulting tour is “long”. A better approximation is possible here using the Combination Lemma, which will give a bound of  $O(\log r)$  instead of  $O(r)$ , but we cannot completely remove  $r$  from the approximation factor.

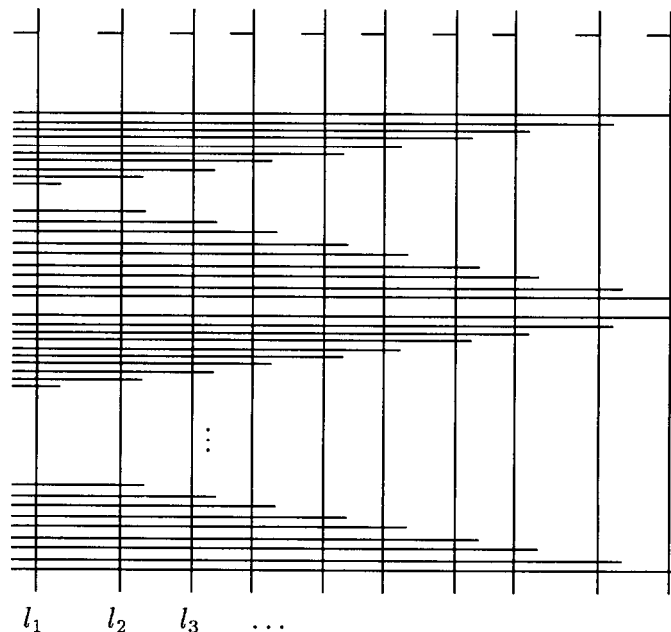


Fig. 7.  $r$  appears in ratio of  $APX/OPT$ .

## 5. Counterexamples

This section contains examples that show that some other (more naive) methods for picking a representative point in each region may not yield a constant error ratio. The first such example is the most naive. Pick an arbitrary point in each region. It is easy to generate examples that show that even when all regions are unit segments parallel to the  $x$ -axis, and the representing point is chosen to be the middle point of the segment, the result can be arbitrarily bad. Start with an arbitrary simple polygon in the plane, whose perimeter is much longer than its height. We make this polygon an optimal tour on the midpoints of some segments, by placing many (equal-length horizontal) segment midpoints along it. However, if the segments are long enough, such that one vertical line suffices to stab all segments, then an optimal tour will be a vertical line segment of length twice the height of the polygon, which is small relative to its perimeter.

Other possible choices for representing points are based on trying to extend the tree heuristic which works so well for point regions (i.e., the classical Euclidean TSP). In this heuristic we build a minimum spanning tree on the points, complete it to an Euclidean tour which is then shortcut to a TSP tour. There are several ways in which we can think of building a minimum spanning tree on a set of parallel unit segments (which are all equivalent for points in the plane). Suppose we have somehow connected a subset of the segments into a forest, and picked representing points on them. We can pick the next segment to be connected by a “Prim” type algorithm: Pick a point on some segment not yet visited, that is closest to points already selected. Alternatively we can think of a “Kruskal” type algorithm that decides next to connect any two segments whose distance between them is minimized, as long as no cycles are closed. Of course for parallel segments, the points on the segments minimizing the distance may not be unique, so this algorithm is not fully specified. There are two alternatives: The simple alternative is to choose one such minimizing pair arbitrarily, the second, and more complex is to try to optimize over all choices of minimizing points. Unfortunately, we do not know how to accomplish this in polynomial time, so we do not analyze the possible success of this second method (and we refer to the first alternative as our “Kruskal” type algorithm).

Intuitively it is not surprising that both algorithms fail to produce the desired result, as the examples of Figs. 8 and 9 show. The Prim type algorithm picks a representative point in a myopic fashion, a choice that may prove to be disastrous later. The Kruskal type algorithm fails for the opposite reason of allowing too much flexibility in the choice of which point(s) will be used to connect the segments and thus allows the salesman to travel within the regions “for free”.

The approximation in Fig. 8 (shown by a dotted line) picks the nearest point on a segment not yet connected, and is thus made to zigzag, while the optimal tour is a rectangle of much shorter length. Notice that all the segments in this example can be made to be of equal length by extending them to the right or left appropriately.

In the example of Fig. 9, each of the middle segments has two points chosen by a “Kruskal” type algorithm. If we build our approximation by connecting points

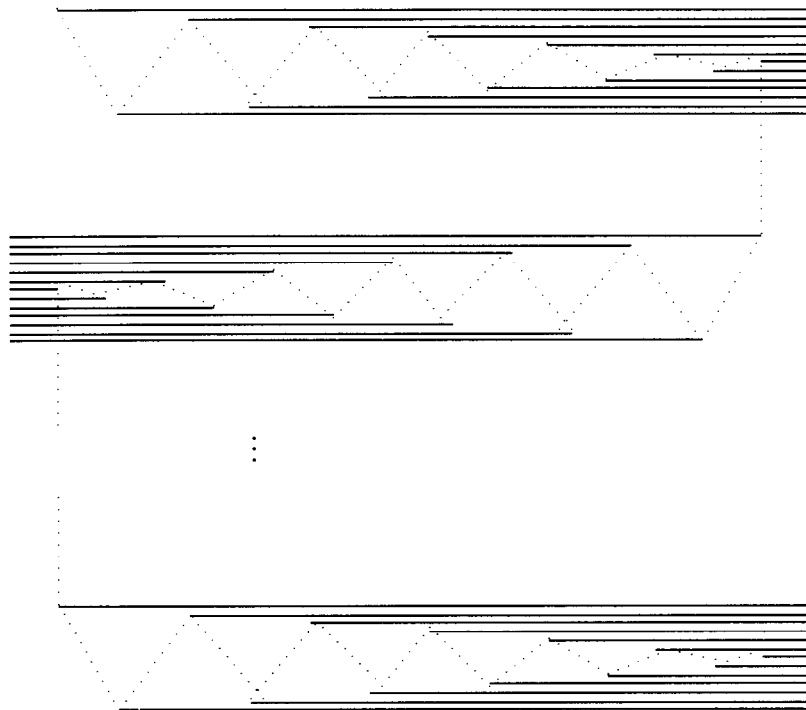


Fig. 8. Failure of Prim type algorithm.

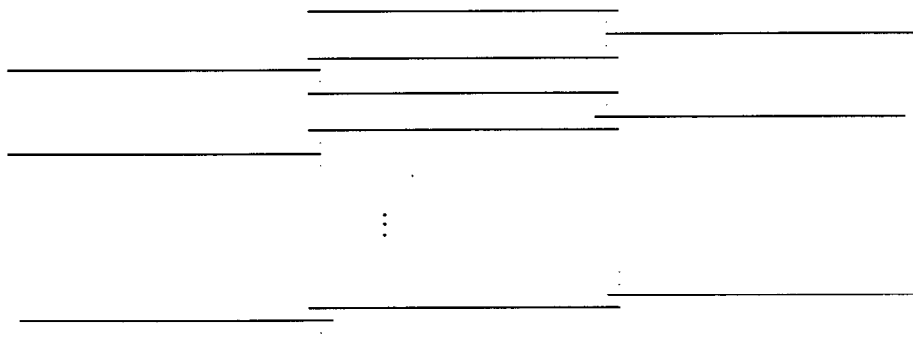


Fig. 9. Failure of Kruskal type algorithm.

chosen in the same region by the segment they share a very poor approximation results. Alternatively, we can choose to include all points selected by the Kruskal type algorithm. as input to a approximation TSP algorithm, but this too (although successful in the example given) may yield bad approximations in general.

## 6. Concluding remarks

We conclude this paper by mentioning some open problems. These correspond to cases for which we have not been able to find a polynomial time algorithm to approximate, with a bounded performance guarantee, an optimal TSP tour (or prove that no such approximation exists unless  $P = NP$ ).

The first such open problem is for regions that are non-uniform parallel elements. One would like a bound independent of the ratio  $r$  between longest and shortest segment. We assume that the number of distinct sizes is not fixed, and more strongly, that the segments cannot be divided into classes, such that within each class the ratio of the longest to shortest segment is small. Otherwise an approximation can be found by combining approximations for the individual classes.

The second open problem concerns non-parallel unit segments (where the number of directions is not fixed, and the ratio of their projections on a given direction is not bounded).

The third open problem concerns regions (convex or non-convex) which can be quite general, as long as their diameter is known. Here we cannot apply our minimum perimeter rectangle approach, because we may not be able to efficiently compute such a rectangle.

A related question is whether we can approximate non-connected regions, such as regions each of which is comprised of two points.

Other generalizations may be quite straightforward, such as regions in higher dimensions.

## References

- [1] A. Aggarwal and S. Suri, Fast algorithms for computing the largest empty rectangle, in: *Proceedings of the 3rd Annual ACM Symposium on Computational Geometry* (1987) 278–289.
- [2] S. Anily and R. Hassin, The swapping problem, *Networks* 22 (1992) 419–433.
- [3] E.M. Arkin, S.P. Fekete and J.S.B. Mitchell, The lawnmower problem, submitted.
- [4] E. Benavent, V. Campos, A. Corberan and E. Mota, Analisis de heurísticos para el problema del cartero rural, *Trab. Estadística Invest. Oper.* 36 (1985) 27–38.
- [5] D. Bienstock, M. Goemans, D. Simchi-Levi and D. Williamson, A note on the prize collecting traveling salesman problem, Technical Report, Columbia University (1990).
- [6] N. Christofides, Worst-case analysis of a new heuristic for the traveling salesman problem, Technical Report, GSIA, Carnegie-Mellon University (1976).
- [7] G. Cornuejols and G.L. Nemhauser, Tight bounds for Christofides' traveling salesman heuristic, *Math. Programming* 14 (1978) 116–121.

- [8] J.T. Current and D.A. Schilling, The covering salesman problem, *Transportation Sci.* 23 (1989) 208–213.
- [9] G.N. Frederickson, Approximation algorithms for some postman problems, *J. ACM* 26 (1979) 208–213.
- [10] G.N. Frederickson, M.S. Hecht and C.E. Kim, Approximation algorithms for some routing problems, *SIAM J. Comput.* 7 (1978) 178–193.
- [11] A.M. Frieze, Worst-case analysis of algorithms for the traveling salesman problem, *Oper. Res. Verfahren* 32 (1979) 93–112.
- [12] A.M. Frieze, An extension of Christofides heuristic to the  $k$ -person traveling salesman problem, *Discrete Appl. Math.* 6 (1983) 79–83.
- [13] M.R. Garey, R.L. Graham and D.S. Johnson, Some NP-complete geometric problems, in: *Proceedings of the 8th Annual ACM Symposium on Theory of Computing* (1976) 10–22.
- [14] D.S. Johnson and C.H. Papadimitriou, Performance guarantees for heuristics, in: E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan and D.B. Shmoys, eds., *The Traveling Salesman Problem*, Ch. 5 (Wiley, New York, 1985).
- [15] S. Ntafos, Watchman routes under limited visibility, *Comput. Geom.* 1 (1992) 149–170.
- [16] C.H. Papadimitriou, The Euclidean traveling salesman problem is NP-complete, *Theoret. Comput. Sci.* 4 (1977) 237–244.
- [17] D.J. Rosenkrantz, R.E. Stearns and P.M. Lewis II, An analysis of several heuristics for the traveling salesman problem, *SIAM J. Comput.* 6 (1977) 563–581.